CLAIMS

1. A method for operating a first server to receive backup service from a second server when the first server detects an operational fault and decides to discontinue handling service requests, the method comprising the steps of:

sending a first indication from the first server to the second server when the first server detects the operational fault that will require it to shut down;

sending a second indication from the first server to the second server indicating the type of operational fault detected by the first server;

receiving a shutdown command at the first server from the second server if the second server can provide backup service to the first server;

completing existing service requests to the first server; and

shutting down the first server.

2. The method in accordance with claim 1 further comprising the step of sending a periodic request from the second server to the first server to stay shut down, after the first server has shut down and its services are being provided by the second server.

3. The method in accordance with claim 2 further comprising the step of rebooting the first server after the operational fault has been cured.

4. The method in accordance with claim 1 further comprising the step of monitoring the status of the first server to detect any operational faults prior to sending of the first indication to the second server in the event an operational fault is detected.

5. The method in accordance with claim 1 further comprising the step of letting the first server complete certain functions it is performing at the time the operational fault is detected before it shuts down.

6. The method in accordance with claim 5 further comprising the step of sending a periodic request from the second server to the first server to stay shut down, after the first server has shut down.

7. A method for operating a first server to provide backup service to a second server when the first server detects an operational fault, the method comprising the steps of:

receiving a first request at the second server requesting the second server to take over the operations of the first server when the first server has detected an operational fault that will require it to shut down;

permitting the first server to complete certain functions it is performing at the time the operational fault is detected before it shuts down;

taking over the functions of the first server by the second server after the first server has shut down.

8. The method in accordance with claim 7 further comprising the step of monitoring the operational status of the first server to detect any operational faults and cause sending of the first request to the second server to take over the operations of the first server in the event an operational fault is detected on the first server.

9. The method in accordance with claim 8 further comprising the steps of:

determining if the second server can provide backup service to the first server; and

requesting the first server to shut down if the second server can provide backup service to the first server.

10. The method in accordance with claim 9 further comprising the step of receiving an indication at the second server from the first server indicating the type of operational fault detected by the first server.

11. The method in accordance with claim 10 further comprising the step of receiving a periodic request at the first the first server sent from the second server for the first server to stay shut down.

12. The method in accordance with claim 11 further comprising the step of rebooting the first server after the detected operational fault has been cured.

13. Apparatus for transferring service requests from a first server to a second server in a cluster of servers when the first server has a fault therein, said apparatus comprising:

first means in said first server for monitoring its operations, said first monitoring means generating a first fault signal when it is determined that said first server has a fault therein, said first fault signal being sent to said second server; and

wherein when said second server receives said first fault signal it permits said first server to complete existing service request that are being processed by said first server and then takes over and processes further service requests directed to said first server.

14. Apparatus for transferring service requests from a first server to a second server in a cluster of servers when the first server has a fault therein, and said first server sends a fault signal to said second server upon there being a fault in said first server, said apparatus comprising:

first means in said second server for receiving said fault signal from said first server; and

wherein when said second server receives said first fault signal it permits said first server to complete existing service request that are being processed by said first server and then takes over said first server and processes further service requests directed thereto.

15. The apparatus in accordance with claim 14 further comprising memory means in said second server, said first server generates a compilation of the status of service requests being processed by said first server and sends said compilation to said second server to be stored in said memory means thereat, said compilation being used by said second server to take over processing of service requests directed to said first server in an orderly fashion.

16. The apparatus in accordance with claim 15 wherein after said second server takes over processing of service requests directed to said first server, said second server transmits stay dead signals to said first server causing it not to attempt to commence processing new service requests.

17. The apparatus in accordance with claim 16 wherein said first and said second servers are file servers.

18. A computer readable medium containing executable instructions for providing backup service between a first server and a second server operating in a cluster mode when the first

servers experiences a fault that will cause it to shut down, the executable program instructions comprising program instructions for:

sending a fault signal from the first server to the second server;

determining if the second server can provide backup service to the first server;

requesting the first server to shut down if the second server can provide backup service to the first server; and

taking over the functions of the first server by the second server after the first server has complete current operations and has shut down.

19. Apparatus for operating a first server to receive backup service from a second server when the first server detects an operational fault, said apparatus comprising:

means for sending a first indication from the first server to the second server when the first server detects the operational fault that will require it to shut down;

means for sending a second indication from the first server to the second server indicating the type of operational fault detected by the first server;

means for receiving a shutdown command from the second server at the first server if the second server can provide backup service to the first server; and

means for shutting down the first server.

20. Apparatus for operating a first server to provide backup service to a second server when the first server detects an operational fault, said apparatus comprising:

means for receiving a first request at the second server from the first server requesting the second server to take over the operations of the first server when the first server has detected an operational fault that will require it to shut down;

means for permitting the first server to complete certain functions it is performing at the time the operational fault is detected before it shuts down; and

means for taking over the functions of the first server at the second server after the first server has shut down.

*****